الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي و البحث العلمي
ⵓⵖⵏⵉⵣ ⵏ ⵓⵙⵏⵎⵙⵙⴰ ⵙⵏⵜⵉⵣ ⴰ ⵙⵏⴰⵍⵉ ⵙⵓⵓⵏⴰ

مرجع التكوين الأوّلي في الطور الثالث

" تقنيات وأدوات الذكاء الاصطناعي "

2024-2025

اللجنة الوطنية للإشراف ومتابعة تنفيذ برنامج التكوين في الطور الثالث بمؤسسات التعليم العالي:

▪ أ. بن يمينة سعيد	رئيس اللجنة الوطنية
▪ أ. رشيدة بوعلوش	مديرة التكوين في الدكتوراه
▪ أ. زبوشي محمد عبد الرؤوف	نائب مدير مكلف بالطور الثالث
▪ أ. بوعروري جعفر	رئيس الندوة الجهوية لجامعات الوسط
▪ أ. لطرش محمد الهادي	رئيس الندوة الجهوية لجامعات الشرق
▪ أ. شعلال أحمد	رئيس الندوة الجهوية لجامعات الغرب

لجنة خبراء مادة تقنيات وأدوات الذكاء الاصطناعي:

▪ أ.د. عبد الوهاب موساوي	رئيس اللجنة
▪ أ.د. بغداد عثماني	خبير
▪ أ. د. صابر بن حرز الله	خبير
▪ أ. د. شوقي جدي	خبير
▪ د. سليمان بالاعور	خبير
▪ د. عطية نهار	خبير
▪ د. عمر طالبي	خبير
▪ د. عبد الحكيم شريط	خبير
▪ أ.د. عبد الرحمان يوسفات	خبير

جدول المحتويات

1 المقدمة ____ 4

1.1 سياق وأهمية الذكاء الاصطناعي ____ 4

2.1 رهانات وأهداف التكوين ____ 4

3.1 دمج أدوات الذكاء الاصطناعي في تقنيات المعلومات والاتصال والبحث ____ 4

2 أهداف التكوين ____ 5

1.2 التعرف على المفاهيم الأساسية للذكاء الاصطناعي ____ 5

2.2 إتقان معالجة وتحليل كميات كبيرة من البيانات باستخدام أدوات الذكاء الاصطناعي ____ 5

3.2 إتقان عملية تطوير نماذج الذكاء الاصطناعي ____ 5

4.2 فهم و إتقان أدوات الذكاء الاصطناعي لتصور البيانات ____ 5

5.2 القدرة على تطبيق أدوات الذكاء الاصطناعي لحل مشكلات بحثية ____ 5

6.2 ضمان استخدام مسؤول وموجه للذكاء الاصطناعي والبيانات ____ 5

3 تنظيم التكوين ____ 6

1.3 المتطلبات الأساسية ____ 6

2.3 فريق التكوين ____ 6

3.3 تكوين المكونين ____ 6

4.3 الوسائل المادية والبرمجية ____ 6

5.3 الأنشطة البيداغوجية المقترحة لمادة "تقنيات وأدوات الذكاء الاصطناعي" ____ 6

6.3 منهجية التنظيم، المحتويات التعليمية، وآليات التقييم ____ 8

7.3 برنامج المحاضرات المقترحة: ____ 9

8.3 برنامج الورشات المقترحة ____ 10

4 البرنامج المفصل ____ 11

المحور الأول: دور واستخدام الأدوات الحرة والمفتوحة المصدر في الذكاء الاصطناعي والبرمجة (8 ساعات) ____ 11

المحور الثاني: التعلم والتواصل المعززان بالذكاء الاصطناعي (8 ساعات) ____ 13

المحور الثالث: البحث وتحليل البيانات المدعوم بالذكاء الاصطناعي (24 ساعة) ____ 16

المحور الرابع: توظيف الذكاء الاصطناعي في حل المشكلات البحثية (8 ساعات) ____ 27

5 المراجع العلمية ____ 30

1. المقدمة

1.1. سياق وأهمية الذكاء الاصطناعي

أصبح الذكاء الاصطناعي اليوم محور التحول الرقمي والمجتمعي. وله تأثير على العديد من القطاعات، بما في ذلك الصحة والمالية والصناعة والعلوم الاجتماعية والبحث العلمي. من خلال دمج الذكاء الاصطناعي في العمليات الأكاديمية والمهنية، يمكن أتمتة المهام المعقدة، وتحسين تحليل البيانات الضخمة وتحسين عملية صنع القرار. في البحث العلمي، يوفر الذكاء الاصطناعي آفاقًا جديدة لاستكشاف ونمذجة الظواهر المعقدة.

في الأوساط الجامعية، بات إتقان تقنيات وأدوات الذكاء الاصطناعي مهارة لا غنى عنها. ويحتاج طلبة الدكتوراه إلى أن يكونوا قادرين على استخدام هذه الأدوات لإنجاز بحوث متقدمة، وتحليل بيانات كبيرة وكتابة منشورات علمية عالية الجودة. يهدف هذا التكوين إلى توفير أرضية صلبة في الذكاء الاصطناعي مع تعزيز نهج نقدي وأخلاقي في استخدامه.

2.1. رهانات وأهداف التكوين

يطرح الاستخدام المتزايد للتقنيات القائمة على الذكاء الاصطناعي العديد من التحديات، خاصة فيما يتعلق بالشفافية، والتحيزات الخوارزمية، وحماية البيانات. لذلك من الضروري تكوين الباحثين في المبادئ الأساسية للذكاء الاصطناعي وتطبيقاته وتداعياته الأخلاقية.

يهدف هذا التكوين إلى:

- **فهم دور الأدوات مفتوحة المصدر واستخدامها في الذكاء الاصطناعي:** تحديد مزايا الحلول مفتوحة المصدر لتطوير الذكاء الاصطناعي وتطبيقه، وإتقان استغلالها في سياق أكاديمي وعلمي.
- **اكتساب إتقان مفاهيم الذكاء الاصطناعي:** فهم الأسس النظرية والعملية للتعلم الآلي والتعلم العميق وعلم البيانات.
- **تطوير القدرة على استعمال أدوات الذكاء الاصطناعي المتقدمة:** استخدام أطر عمل مثل TensorFlow و PyTorch و Scikit-learn لتصميم النماذج الذكية وتدريبها وتقييمها.
- **تطبيق الذكاء الاصطناعي في البحث العلمي:** استغلال أدوات الذكاء الاصطناعي لتحليل البيانات والنمذجة واستخراج المعرفة في مختلف المجالات الأكاديمية.
- **تعزيز المهارات في معالجة البيانات وعرضها:** استخدام مكتبات مثل Pandas وMatplotlib وSeaborn لاستكشاف البيانات وعرضها بفعالية.
- **رفع مستوى وعي طلبة الدكتوراه بالاستخدام المنضبط والمسؤول للذكاء الاصطناعي:** دمج التقدم في التقنيات الجديدة وأفضل الممارسات في هذا المجال مع ضرورة تزويدهم بتكوين صارم يمكّنهم من تكييف الذكاء الاصطناعي مع المتطلبات والاحتياجات المحددة لأبحاثهم الخاصة.

3.1. دمج أدوات الذكاء الاصطناعي في تقنيات المعلومات والاتصال والبحث

يعزز الذكاء الاصطناعي أدوات تكنولوجيا المعلومات والاتصالات من خلال توفير حلول للأتمتة والتحسين والمساعدة الذكية. واليوم، تسهل العديد من أدوات الذكاء الاصطناعي البحث والكتابة الأكاديمية:

- المساعدة في الكتابة والتصحيح: Grammarly, LanguageTool, DeepL.
- التحقق من الانتحال والنزاهة العلمية: Turnitin، Compilatio، Plagscan.
- إدارة المراجع والببليوغرافيا: Zotero، Mendeley، EndNote.
- تحليل البيانات وعرضها: Power BI, Tableau, Matplotlib, Seaborn.
- الأدوات التعاونية والأكاديمية: Google Scholar، ResearchGate، Overleaf، GitHub.

تعتبر هذه الأدوات جزءا من بيئة بحثية رقمية متزايدة التطور، تتيح لطلبة الدكتوراه أن يكونوا أكثر كفاءة ودقة في أعمالهم العلمية.

2. أهداف التكوين

1.2. التعرف على المفاهيم الأساسية للذكاء الاصطناعي

- التعرف على المبادئ الأساسية للذكاء الاصطناعي، بما في ذلك التعلم المُوجَّه، التعلم غير المُوجَّه، والتعلم المعزز.
- فهم الفرق بين الذكاء الاصطناعي، التعلم الآلي، والتعلم العميق.
- تحديد مجالات تطبيق الذكاء الاصطناعي في البحث العلمي والصناعة.

2.2. إتقان معالجة وتحليل كميات كبيرة من البيانات باستخدام أدوات الذكاء الاصطناعي

- تعلم كيفية جمع وتنقية وهيكلة البيانات لاستخدامها بواسطة خوارزميات الذكاء الاصطناعي.
- استخدام أدوات مثل Pandas و Numpy لمعالجة مجموعات كبيرة من البيانات.
- تطبيق تقنيات تقليل الأبعاد وتحويل البيانات.

3.2. إتقان عملية تطوير نماذج الذكاء الاصطناعي

- تصميم وتنفيذ نماذج الذكاء الاصطناعي باستخدام مكتبات مثل R وTensorFlow وPyTorch وScikit-learn.
- فهم مراحل تدريب النماذج، التحقق من صحتها، واختبارها.
- تحسين أداء النماذج من خلال ضبط معلمات الضبط (Hyperparameters) واستخدام تقنيات التحقق المتقاطع (Cross-Validation).

4.2. فهم و إتقان أدوات الذكاء الاصطناعي لتصور البيانات

- إتقان استخدام مكتبات مثل Matplotlib وSeaborn وPower BI لعرض البيانات.
- تعلم كيفية تفسير النتائج من خلال تصورات تفاعلية.
- إدارة لوحات التحكم التحليلية (Dashboards) لعرض استنتاجات قابلة للتنفيذ.

5.2. القدرة على تطبيق أدوات الذكاء الاصطناعي لحل مشكلات بحثية

- تحديد حالات استخدام الذكاء الاصطناعي في البحث العلمي (الصحة، البيولوجيا، العلوم الاجتماعية، الهندسة، إلخ).
- تجربة النماذج المدربة مسبقاً لتسريع تطوير حلول مبتكرة.
- تطبيق تقنيات معالجة اللغة الطبيعية (NLP) والرؤية الحاسوبية (Computer Vision) لحل مشكلات محددة.

6.2. ضمان استخدام مسؤول وموجه للذكاء الاصطناعي والبيانات

- استيعاب التحديات المرتبطة باستخدام الذكاء الاصطناعي وعلم البيانات في السياقات المهنية والعلمية.
- تعلم كيفية التعرف على التحيزات الخوارزمية وتجنبها لضمان نتائج موثوقة وعادلة.
- الامتثال للوائح السارية بشأن حماية البيانات والاستخدام المسؤول للذكاء الاصطناعي.

3. تنظيم التكوين

1.3. المتطلبات الأساسية

يجب أن يمتلك طلبة الدكتوراه أساسيات في البرمجة (Python)، الإحصائيات، ومعالجة البيانات. يفضل أن تكون لديهم معرفة متقدمة بالبرمجيات المكتبية وبأساليب البحث الوثائقي.

2.3. فريق التكوين

يُقدم التكوين من طرف:

- **أساتذة باحثين** متخصصين في الذكاء الاصطناعي وعلم البيانات.
- **باحثين في مجال** الذكاء الاصطناعي يتمتعون بخبرة في تطبيقات الذكاء الاصطناعي في البحث العلمي.

3.3. تكوين المكونين

سيتم تنظيم دورة تكوينية للمكونين لضمان التجانس في تدريس مختلف الوحدات التعليمية.

4.3. الوسائل المادية والبرمجية

لضمان تكوين فعال وتفاعلي، سيستفيد طلبة الدكتوراه من الوسائل والبرمجيات التالية:

1.4.3. الوسائل المادية

- أجهزة حواسيب مزودة بوحدات معالجة رسومية (GPU) لتدريب نماذج الذكاء الاصطناعي.
- اتصال بالإنترنت وقاعات مجهزة للجلسات التطبيقية.

2.4.3. البرمجيات

سيستخدم طلبة الدكتوراه للبرمجة، تحليل البيانات، والنمذجة البرمجيات وبيئات التطوير التالية:

الفئة	الأدوات
لغات البرمجة	Python, R
مكتبات الذكاء الاصطناعي والتعلم الآلي	TensorFlow, PyTorch, Scikit-learn, Keras, Caret (R), FactOminer ®, FactoShiny ®
بيئات التطوير	Jupyter Notebook (Python), RStudio (R)
معالجة البيانات وتحليلها	Pandas, NumPy (Python), dplyr, tidyr (R)
عرض البيانات وتحليلها	Matplotlib, Seaborn (Python), ggplot2, Shiny (R), Power BI, Tableau, Dash

5.3. الأنشطة البيداغوجية المقترحة لمادة "تقنيات وأدوات الذكاء الاصطناعي"

تُنفَّذ الأنشطة البيداغوجية المقترحة بطريقة منهجية عبر مجموعة متنوعة من المحاضرات، الندوات، وورش العمل، بما يتوافق مع الكفاءات المحددة في الملحق 1 وأهداف البرنامج المدرجة في الملحق 2. ويرتكز البرنامج على أربعة محاور أساسية:

المحور الأول: دوروا استخدام الأدوات الحرة ومفتوحة المصدر في الذكاء الاصطناعي والبرمجة (8 ساعات)

- مقدمة عن البرمجيات الحرة ومفتوحة المصدر (مثل تراخيص GPL وMIT وغيرهما).
- نظرة عامة حول الأدوات مفتوحة المصدر المستخدمة في البحث والتعليم.
- استغلال الأدوات المفتوحة المصدر لإدارة المشاريع وتعزيز العمل التعاوني

جدول – المحور الأول

الرقم	النشاط	المدة	وصف مختصر	المهارات المستهدفة
1	مقدمة إلى مفاهيم البرمجيات الحرة	ساعة واحدة	استكشاف المبادئ الأخلاقية والقانونية لتراخيص البرمجيات مفتوحة المصدر (مثل GPL، MIT، Apache).	م1، م26
2	اكتشاف الأدوات مفتوحة المصدر للتعليم والبحث	ساعة واحدة	التعرف على استخدام أدوات مثل Zotero، LibreOffice، Jupyter، R، Python، وOverleaf.	م13، م26
3	تطبيق عملي باستخدام أدوات مفتوحة المصدر	ساعتان ونصف	اكتساب الخبرة في استخدام أدوات مثل Jupyter (للتحليل)، Zotero (لإدارة المراجع)، وLibreOffice (للتحرير).	م13، م 21
4	إدارة تعاونية للمشاريع العلمية المفتوحة	ساعتان ونصف	تعلم استخدام Git/GitLab وOverleaf لإدارة المشاريع التعاونية والتحرير المفتوح.	م13، م21
	المجموع العام	8 ساعات		

المحور الثاني: التعلم والتواصل المعززان بالذكاء الاصطناعي

- مقدمة حول الذكاء الاصطناعي وتأثيراته على التعلم والتواصل.
- دعم الكتابة العلمية باستخدام تقنيات الذكاء الاصطناعي.
- البحث الوثائقي والتحليل النقدي المدعومان بالذكاء الاصطناعي.

جدول – المحور الثاني

الرقم	النشاط	المدة	وصف مختصر	المهارات المستهدفة
1	التوعية بأساسيات الذكاء الاصطناعي	ساعتان	فهم المبادئ الأساسية للذكاء الاصطناعي وتأثيراته في مجال التواصل العلمي.	م1، م2، م4
2	تعزيز القدرات التحريرية باستخدام أدوات ذكية	3 ساعات	الاستفادة من أدوات مثل Grammarly، LanguageTool، ChatGPT، DeepL وDeepSeek لتحسين جودة الكتابة العلمية.	م11، م12، م13
3	تعزيز البحث المنهجي باستخدام أدوات ذكية للبحث الوثائقي	3 ساعات	استخدام أدوات مثل Elicit، Semantic Scholar، Scite وManus لتحسين عمليات البحث والتحليل النقدي للوثائق الأكاديمية.	م4، م13، م17
	المجموع العام	8 ساعات		

المحور الثالث: البحث وتحليل البيانات المدفوع بالذكاء الاصطناعي باستخدام Python

- مقدمة في علم البيانات باستخدام Python.
- المعالجة الأولية وتنقية البيانات.
- التجميع، الانضمام، وتجميع البيانات.
- التحليل الاستكشافي والإحصائيات.
- عرض البيانات.
- النمذجة التنبؤية وتقنيات التعلم العميق.
- إنشاء لوحة تحكم تفاعلية.

جدول – المحور الثالث

الرقم	النشاط	المدة	وصف مختصر	المهارات المستهدفة
1	أساسيات علم البيانات	ساعتان	مقدمة إلى Python وPandas وبيئة علم البيانات	م16، م20
2	تنقية وهيكلة البيانات	3 ساعات	إدارة وتصحيح البيانات المفقودة والكشف عن القيم الشاذة	م19، م23
3	العمليات المتقدمة على البيانات	ساعتان	إتقان العمليات المتقدمة مثلgroupby وmerge وagg	م20، م23
4	التحليل الإحصائي الاستكشافي	ساعتان	حساب المتوسطات، التباينات، وإجراء اختبارات ANOVA	م23
5	تقنيات العرض البياني المتقدم	3 ساعات	استخدام مكتبات مثل Seaborn و Plotly	م23، م25
6	تقنيات التنبؤ المتقدمة والتعلم العميق	10 ساعات	استكشاف النماذج التقليدية لتعلم الآلة والتعرف على الشبكات العصبية باستخدام Keras	م19، م22، م24
7	تطوير تطبيقات تفاعلية	ساعتان	إنشاء لوحات تحكم تفاعلية باستخدام أدوات مثلStreamlit وDash	م24، م25
	المجموع العام	24 ساعة		

المحور الرابع: استخدام الذكاء الاصطناعي لحل مشاكل بحث

- إنشاء لوحات تحكم تفاعلية لتحليل البيانات.
- دراسات حالة: تطبيق الذكاء الاصطناعي في مجالات متنوعة.
- التجريب على قواعد بيانات علمية مختلفة.

جدول – المحور الرابع

الرقم	النشاط	المدة	وصف مختصر	المهارات المستهدفة
1	عرض تفاعلي متقدم لنتائج علمية	ساعتان	تطوير واجهات تفاعلية لتحليل أفضل للبيانات العلمية باستخدام Power BI، Tableau، Shiny	م25
2	حلول ملموسة لمشاكل علمية بواسطة الذكاء الاصطناعي	3 ساعات	تنفيذ حلول تعتمد على الذكاء الاصطناعي لمشاكل متخصصة في مجالات مختلفة مثل OCR، NLP، CNN	م19، م20، م21، م22، م23، م24
3	استكشاف تحليلي متقدم لقواعد بيانات علمية	3 ساعات	التعامل مع قواعد بيانات حقيقية، مع التحضير والتحليل، باستخدام مجموعات بيانات متوفرة على Kaggle أوUCI	م19، م23، م24
	المجموع العام	8 ساعات		

6.3. منهجية التنظيم، المحتويات التعليمية، وآليات التقييم

تهدف مادة تقنيات وأدوات الذكاء الاصطناعي إلى تمكين طلبة الدكتوراه من اكتساب المهارات النظرية والتطبيقية اللازمة لاستخدام الأدوات والتقنيات المتقدمة للذكاء الاصطناعي بكفاءة في أبحاثهم العلمية.

1.6.3. أساليب التقييم وتنظيم التكوين: يتضمن التكوين نظامين متكاملين للتقييم، يهدفان إلى ضمان اكتساب الكفاءات المستهدفة بشكل فعّال:

- **التقييم التكويني (من خلال المحاضرات):** يتم هذا التقييم أثناء المحاضرات، حيث يساعد على متابعة مستوى فهم طلبة الدكتوراه للمفاهيم النظرية بشكل تدريجي. كما يتيح الكشف عن التحديات التي يواجهها الطلبة والعمل على تعديل الأساليب التعليمية بشكل فوري لتعزيز عملية التعلم.
- **التقييم التراكمي (عبر الورشات التطبيقية والمشاريع العلمية):** يتم هذا التقييم أثناء الورشات والمشاريع العملية، بهدف قياس مدى اكتساب الطلبة للكفاءات المطلوبة عمليًا. ويركز التقييم على قدرة الطلبة على معالجة المشكلات العلمية الواقعية، توظيف المعارف النظرية في مواقف عملية والاستخدام الفعّال لأدوات وتقنيات الذكاء الاصطناعي في السياقات التطبيقية.

2.6.3. **نوع التكوين المقترح:** محاضرات، ورش عمل تطبيقية، ومشاريع علمية.

3.6.3. **الحجم الساعي الإجمالي:** 48 ساعة.

- 20 ساعة نظرية على شكل محاضرات.
- 28 ساعة على شكل ورش عمل تطبيقية ومشاريع علمية).

جدول تفصيلي للحجم الساعي لكل سداسي

السداسي	الأنشطة التعليمية	إجمالي الحجم الساعي	التوزيع التفصيلي
الأول	المحاضرات	8 ساعات	من المحاضرة 1 إلى المحاضرة 4
	ورش العمل التطبيقية	8 ساعات	من ورشة العمل 1 إلى ورشة العمل 3
الثاني	المحاضرات	12 ساعة	من المحاضرة 5 إلى المحاضرة 7
	ورش العمل التطبيقية	20 ساعة	من ورشة العمل 4 إلى ورشة العمل 8
إجمالي الحجم الساعي	**المجموع السنوي**	48 ساعة	**نظري: 20 ساعة** **تطبيقي: 28 ساعة**

7.3. برنامج المحاضرات المقترحة:

تقيّم المحاضرات تقييما تكوينيا، حيث تُساعد طلبة الدكتوراه على تطوير فهم عميق ومتين للأسس النظرية والمنهجيات الأساسية لتطبيق أدوات الذكاء الاصطناعي بفعالية.

المهارات المستهدفة	عنوان المحاضرة	أهداف المحاضرة	محتوى المحاضرة	التحديات العلمية المعالجة	المدة الزمنية	نوع التقييم
م1، م13، م26	**محاضرة 1:** أدوات المصادر المفتوحة في البحث العلمي	إتقان أدوات المصادر المفتوحة الأساسية للبحث العلمي	التراخيص (GPL، MIT)، Zotero، Jupyter، Overleaf	التحديات المرتبطة بإعادة إنتاج النتائج العلمية	ساعتان	تكويني
م1، م2، م4	**محاضرة 2:** أساسيات الذكاء الاصطناعي وآثاره العلمية	فهم شامل للمفاهيم الأساسية للذكاء الاصطناعي في إطار علمي	التعلم الآلي والآثار العلمية الحقيقية	دمج الذكاء الاصطناعي في التعليم والتواصل العلمي	ساعتان	تكويني
م11، م12، م13	**محاضرة 3:** الكتابة العلمية باستخدام الذكاء الاصطناعي	تعزيز مهارات الكتابة العلمية باستخدام الذكاء الاصطناعي	أدوات مثل Grammarly، LanguageTool، DeepL ،ChatGPT	تحسين جودة الكتابة اللغوية والعلمية في المنشورات	ساعتان	تكويني
م 4، م13، م17	**محاضرة 4:** البحث الوثائقي المتقدم بمساعدة الذكاء الاصطناعي	استخدام الذكاء الاصطناعي لتحسين التحليل النقدي للوثائق	أدوات مثلElicit ، Semantic Scholar، Scite	الصعوبات في الوصول إلى الموارد العلمية واختيارها بشكل ملائم	ساعتان	تكويني
م19، م20، م23	**محاضرة 5:** تحليل إحصائي واستكشافي باستخدام Python	تطوير مهارات قوية في التحليل الإحصائي باستخدام Python	مكتبات Python مثل (Pandas، NumPy)، الإحصائيات الوصفية والاستكشافية	المعالجة الإحصائية لبيانات تجريبية معقدة	4 ساعات	تكويني
م22، م24، م25	**محاضرة 6:** النمذجة التنبؤية والعرض العلمي	تصميم نماذج بحثية تنبؤية وعرضها بكفاءة	التعلم الآلي (Scikit-learn)، التعلم العميق (Keras)، العرض (Plotly، Dash)	الصعوبات في نمذجة البيانات المعقدة في سياق علمي	4 ساعات	تكويني
م24، م25	**محاضرة 7:** مقدمة في التعلم العميق والنمذجة المتقدمة	فهم الشبكات العصبية البسيطة لنمذجة الظواهر المعقدة	التعلم العميق باستخدام Keras (شبكات MLP الأساسية)، التجهيز الزائد، وعرض الطبقات	لماذا وكيف يتم استخدام الشبكات العصبية في البحث التطبيقي؟	4 ساعات	تكويني
إجمالي الحجم الساعي					**20 ساعة**	**تكويني**

8.3. برنامج الورشات المقترحة

تقيم ورش العمل التطبيقية بطريقة تراكمية، وتهدف إلى التحقق العملي من الكفاءات التي اكتسبها طلبة الدكتوراه من خلال التطبيق المباشر للأدوات والتقنيات الخاصة بالذكاء الاصطناعي في سياقات واقعية.

المهارات المستهدفة	عنوان الورشة	أهداف المحاضرة	محتوى المحاضرة	التحديات العلمية المعالجة	المدة الزمنية	نوع التقييم
م13، م21، م26	**ورشة 1:** الاستخدام العملي للأدوات مفتوحة المصدر	إتقان Zotero، Overleaf،Jupyter لإدارة الوثائق العلمية بكفاءة	التثبيت والاستخدام العملي للأدوات	صعوبة إدارة وإعادة إنتاج الأعمال العلمية	3 ساعات	تراكمي
م13، م21	**ورشة 2:** العمل التعاوني المفتوح على المشاريع العلمية	إتقان استخدام Git/GitHub وOverleaf للإدارة التعاونية	عمل تعاوني عملي، إدارة الإصدارات	إدارة معقدة للمشاريع العلمية التعاونية	3 ساعات	تراكمي
م11، م12، م13	**ورشة 3:** أدوات الذكاء الاصطناعي في خدمة الباحث	تعزيز القدرة على البحث عن المصادر العلمية واختيارها وتحليلها باستخدام أدوات الذكاء الاصطناعي	أدوات الكتابة والبحث	جودة الصياغة اللغوية في المنشورات العلمية	ساعتان	تراكمي
م19، م23	**ورشة 4:** تحضير البيانات العلمية وتنقيتها	تجهيز البيانات التجريبية بشكل فعال للتحليل	تنقية ومعالجة البيانات باستخدام Pandas وOpenRefine	الصعوبة في إعداد البيانات التجريبية	4 ساعات	تراكمي
م20، م23	**ورشة 5:** التحليل الإحصائي المطبق على البيانات الحقيقية	إجراء تحليلات إحصائية عملية على البيانات العلمية	ANOVA، الارتباطات، الاختبارات الإحصائية التطبيقية	تعقيدات التحليل الإحصائي في البحث	4 ساعات	تراكمي
م22، م24	**ورشة 6:** التنفيذ العملي للنماذج التنبؤية (التعلم الآلي والعميق)	بناء وتقييم نماذج تنبؤية لحل مشكلات علمية عملية	تطبيق باستخدام Scikit-learn و Keras	التصنيف والتنبؤ بالنتائج العلمية	4 ساعات	تراكمي
م24، م25	**ورشة 7:** تصميم لوحات تحكم تفاعلية	تطوير عروض ديناميكية للأبحاث العلمية	إنشاء لوحات تحكم باستخدام Dash وStreamlit	صعوبة عرض النتائج العلمية المعقدة بوضوح	4 ساعات	تراكمي
م19، م20، م21، م22، م23، م24	**ورشة 8:** مشروع علمي تجريبي شامل باستخدام الذكاء الاصطناعي	تنفيذ مشروع متكامل لتحليل البيانات، نمذجتها وتفسيرها	تطبيق كامل (معالجة أولية، تحليل، نمذجة، وتقييم) على بيانات واقعية	صعوبة استخراج وتثمين المعارف من قواعد البيانات الكبيرة	4 ساعات	تراكمي
إجمالي الحجم الساعي					**28 ساعة**	**تراكمي**

4. البرنامج المفصل

المحور الأول: دور واستخدام الأدوات الحرة والمفتوحة المصدر في الذكاء الاصطناعي والبرمجة (8 ساعات)

يعد استخدام الأدوات المفتوحة المصدر ضروريا لضمان بحث علمي قابل للتكرار، تعاوني، ومتاح للجميع، خاصة في المجالات التي تعتمد على تقنيات الذكاء الاصطناعي. يهدف هذا المحور إلى تمكين طلبة الدكتوراه من التعرف على هذه الأدوات الأساسية وتطوير مهاراتهم في استخدامها بشكل مستقل.

الموضوع الأول: مقدمة عن البرمجيات الحرة ومفتوحة المصدر (مثل تراخيص GPL وMIT وغيرهما).
الموضوع الثاني: نظرة عامة حول الأدوات مفتوحة مصدر المستخدمة في البحث والتعليم.
الموضوع الثالث: استغلال الأدوات المفتوحة المصدر لإدارة المشاريع وتعزيز العمل التعاوني

1- مقدمة عن البرمجيات الحرة ومفتوحة المصدر مثل تراخيص GPL وMIT وغيرهما (ساعتان)

- **الهدف:** فهم المبادئ القانونية والأخلاقية لتراخيص البرمجيات المفتوحة المصدر، بالإضافة إلى تأثيرها على البحث العلمي وتطوير البرمجيات.
- **الأدوات:** لا توجد أدوات محددة مطلوبة، ولكن سيتم تحليل وثائق التراخيص ودراسة مشاريع البرمجيات المفتوحة المصدر مثل GitHub.
- **المحتوى:**
  - مقدمة حول مفهوم البرمجيات الحرة والمفتوحة المصدر.
  - الفروقات بين التراخيص الحرة والتراخيص المملوكة.
  - عرض لأهم التراخيص المفتوحة المصدر:
    - **GPL:** حقوق متروكة، تبادلية.
    - **MIT:** مرنة وبسيطة.
    - **Apache 2.0:** رخصة مرنة تتضمن بندًا خاصًا ببراءات الاختراع.
  - مقارنة بين التراخيص المختلفة (جدول تلخيصي).
  - المسؤوليات القانونية وأفضل الممارسات (مثل الاستشهاد بالمصادر واحترام بنود التراخيص).
  - القضايا الأخلاقية المتعلقة بالنشر الحر للمعرفة.
- **ورشة عمل تطبيقية:**
  - دراسة حالة: تحديد نوع الترخيص المستخدم في أحد المشاريع مفتوحة المصدر على منصة GitHub
  - نقاش: ما هو نوع الترخيص الأنسب للمشاريع الأكاديمية؟

2- نظرة عامة حول الأدوات مفتوحة المصدر المستخدمة في البحث والتعليم (4 ساعات)

الهدف: اكتشاف الأدوات الرئيسية مفتوحة المصدر وفهم فوائدها للبحث والتعليم.

الأدوات: Linux, LibreOffice, Zotero, R, Python, Jupyter Notebook

المحتوى:

- مقدمة في البرمجيات الحرة والمفتوحة المصدر.
  - تاريخ ومبادئ المصدر المفتوح.
  - الفرق بين البرمجيات مفتوحة المصدر والبرمجيات المملوكة.
  - مزايا المصادر المفتوحة للبحث العلمي (الشفافية وقابلية الاستنساخ والتعاون).
- عرض الأدوات مفتوحة المصدر حسب الاستخدام (المكتبية، إدارة المراجع، تحليل البيانات).
- مقارنة مع الأدوات المسجلة الملكية.
- القضايا الأخلاقية والقانونية.
  - التراخيص مفتوحة المصدر (GPL، MIT، Apache، إلخ).
  - المسؤوليات وأفضل الممارسات.
- دور المصادر المفتوحة في استخدام وتعزيز وتطوير الذكاء الاصطناعي.

- **ورشة عمل تطبيقية:**

اختبار أدوات مختلفة وفقًا لاحتياجات طلبة الدكتوراه (مثل تنظيم تنظيم وإدارة المراجع باستخدام Zotero، وتشغيل برنامج نصي من Python على Jupyter).

3- استغلال أدوات مفتوحة المصدر لإدارة المشاريع وتعزيز العمل التعاوني (4 ساعات)

الهدف: استغلال أدوات مفتوحة المصدر لتسهيل إدارة المشاريع والعمل التعاوني في البحث.

الأدوات: Git/GitHub, Nextcloud, Overleaf, Jitsi Meet

المحتوى:

- عرض تقديمي لأدوات الإدارة والتعاون.
- إدارة الإصدارات باستخدام Git والتعاون عبر GitHub.
- المشاركة والتحرير التعاوني للوثائق العلمية.

- **ورشة عمل تطبيقية :**

- العمل كفريق في مشروع بحثي باستخدام GitHub وOverleaf للكتابة التعاونية.

المحور الثاني: التعلم والتواصل المعززان بالذكاء الاصطناعي (8 ساعات)

المحور الثاني من برنامج التكوين يستكشف دور الذكاء الاصطناعي في تعزيز التعلم وتحسين التواصل العلمي، وهو منظم حول ثلاثة محاور متكاملة، يقدم كل منها نهجًا محددًا حول استخدام أدوات الذكاء الاصطناعي في السياق الأكاديمي:

الموضوع الأول: استيعاب المفاهيم الأساسية للذكاء الاصطناعي وتأثيره على التعليم والتواصل.
الموضوع الثاني: استغلال الذكاء الاصطناعي لتحسين الكتابة، والبحث الوثائقي، والتحليل النقدي.
الموضوع الثالث: تطوير نهج نقدي يضمن استخداما فعالا وأخلاقيا لأدوات الذكاء الاصطناعي.
توصية: توفر الأدوات المقترحة غالبا إصدارات مجانية بميزات محدودة. يوصى بالبدء بهذه الإصدارات للحصول على فهم أولي للأدوات. بعد ذلك، يمكن التفكير في الاشتراكات للوصول إلى تحليلات أكثر تقدما. بالإضافة إلى ذلك، يمكن تفضيل البدائل مفتوحة المصدر لما توفره من مرونة أكبر وتحكم أفضل في البيانات.

1- مقدمة عن البرمجيات الحرة ومفتوحة المصدر مثل تراخيص GPL وMIT وغيرهما (ساعتان)

الهدف: فهم أسس الذكاء الاصطناعي وتطوره وتأثيره على التعلم والتواصل.
الأدوات: ChatGPT, Claude, Gemini, Copilot
المحتوى:

- تعريف الذكاء الاصطناعي وتاريخه
  - نشأة وتطور أنظمة الذكاء الاصطناعي.
  - الفرق بين الذكاء الاصطناعي الرمزي، التعلم الآلي والشبكات العصبية.
  - تأثير التطورات الحديثة على التعليم والتواصل.
- مجالات التطبيق
  - في التعليم: الأنظمة التعليمية الذكية، والتصحيح الآلي، والمساعدة في الكتابة.
  - في التواصل: تلخيص النصوص، إعادة الصياغة، والترجمة الآلية.
  - في إنتاج المحتوى :إنشاء المقالات، والتلخيص التلقائي للنصوص.
- القضايا الأخلاقية والتحديات
  - التحيزات الخوارزمية: كيف تؤثر البيانات على النتائج التي تقدمها أنظمة الذكاء الاصطناعي؟
  - الشفافية وقابلية التفسير: لماذا من المهم فهم كيفية عمل أنظمة الذكاء الاصطناعي؟
  - حماية البيانات :ضمان سرية البيانات والامتثال للإطار القانوني المنظم.
  - حماية البيانات والتنظيم :الممارسات الجيدة والإطار القانوني المنظم.

- **ورشة عمل تطبيقية :** تهدف هذه الورشة إلى استكشاف قدرات الذكاء الاصطناعي في تلخيص النصوص الأكاديمية وإعادة صياغتها، مع تطوير رؤية نقدية للنتائج المنتجة من خلال :
  - التعرف على الأدوات: استخدام أدوات مثلChatGPT ، Claude، Gemini، و.Copilot
  - التجريب: إنشاء ملخصات وإعادة صياغة مقاطع أكاديمية.
  - التحليل النقدي: مقارنة النتائج، وتحديد القيود والتحيزات.
  - التطبيق العملي: العمل على وثيقة يختارها طلبة الدكتوراه.
  - المناقشة الجماعية والتقييم التكويني: تبادل الآراء حول التجربة وإجراء التحسينات اللازمة.

2- استغلال الذكاء الاصطناعي لتحسين الكتابة، والبحث الوثائقي، والتحليل النقدي (3 ساعات).

الهدف: تعزيز جودة الكتابة الأكاديمية باستخدام أدوات الذكاء الاصطناعي لتصحيح النصوص، إعادة صياغتها، وتكييفها.

الأدوات: Grammarly، LanguageTool، ChatGPT، DeepL.

المحتوى:

- تصحيح وتحسين النصوص
  - اكتشاف وتصحيح الأخطاء اللغوية (النحوية والإملائية وعلامات الترقيم).
  - تقديم تحسينات للبنية العامة للنص لتعزيز وضوحه وسهولة قراءته.
- تحسين الأسلوب وتكييف النبرة التحريرية
  - إعادة صياغة النصوص العلمية بشكل يعزز من ثرائها اللغوي ودقتها.
  - تعديل مستوى اللغة والنبرة لتتلاءم مع الجمهور المستهدف والمتطلبات الأكاديمية.
- الترجمة وتنسيق النصوص بلغات متعددة
  - الاستفادة من أدوات الذكاء الاصطناعي لترجمة النصوص وتكييفها بما يتناسب مع السياقات المختلفة.
  - مقارنة بين حلول الترجمة المختلفة لتقييم تأثيرها على دقة النصوص وموثوقيتها.

- **ورشة عمل تطبيقية:** تهدف هذه الورشة إلى تطبيق المهارات المكتسبة لتحسين جودة النصوص الأكاديمية باستخدام أدوات الذكاء الاصطناعي.
  - تحليل وتصحيح النصوص تلقائيًا باستخدام Grammarly و.LanguageTool
  - إعادة صياغة النصوص وتحسين الأسلوب الكتابي باستخدام.ChatGPT
  - ترجمة وتكييف مقطع أكاديمي إلى لغات متعددة باستخدام DeepL وبدائل مفتوحة المصدر.
  - مقارنة النتائج بين الأدوات المجانية والمدفوعة.

3- تطوير نهج نقدي يضمن استخداما فعالا وأخلاقيا لأدوات الذكاء الاصطناعي (3 ساعات).

الهدف: يهدف هذا المحور إلى استغلال إمكانيات الذكاء الاصطناعي لتحسين البحث الببليوغرافي وتعزيز التقييم النقدي للمصادر الأكاديمية. يوفر هذا المحور لطلبة الدكتوراه إمكانية الوصول بفعالية أكبر إلى المنشورات ذات الصلة، مع تطوير رؤية نقدية حول موثوقية المعلومات.

الأدوات: Elicit, Semantic Scholar, Scite, Zotero avec IA

المحتوى:

- تحسين البحث الوثائقي:
  - استخدام محركات البحث العلمية المدعومة بالذكاء الاصطناعي لتحديد المقالات ذات الصلة بشكل دقيق.
  - التنقيب في قواعد البيانات الأكاديمية واستخراج المعلومات الرئيسية بطريقة آلية.
- تلخيص المنشورات:
  - إنشاء ملخصات ذكية لتحليل سريع وفعّال للمحتويات الأكاديمية.
  - مقارنة الأساليب اليدوية مع الأساليب الآلية في تقييم المصادر الأكاديمية.
- تحليل الاقتباسات والعلاقات بين المنشورات الأكاديمية لتقييم تأثيرها.
- مراعاة التحيزات الخوارزمية والقيود المتعلقة باستخدام أدوات الذكاء الاصطناعي في البحث العلمي.

- **ورشة عمل تطبيقية :** تهدف الورشة إلى مقارنة البحث الوثائقي التقليدي مع البحث المدعوم بالذكاء الاصطناعي من خلال :
  - إجراء بحث أكاديمي باستخدام Google Scholar والأساليب التقليدية.
  - إجراء بحث آخر باستخدام أدوات مثل Elicit، وSemantic Scholar، وScite، لقياس كفاءة الذكاء الاصطناعي وحدوده في تحليل الوثائق.
  - مناقشة التحيزات الخوارزمية والجوانب الأخلاقية في استخدام هذه الأدوات.

المحور الثالث: البحث وتحليل البيانات المدعوم بالذكاء الاصطناعي (24 ساعة)

يركز المحور الثالث على عملية جمع البيانات ومعالجتها وتحليلها باستخدام أدوات الذكاء الاصطناعي، مع التركيز على إتقان الخطوات الأساسية للنمذجة وتفسير النتائج باستخدام الأساليب الرياضية والخوارزمية الحديثة.

1- مقدمة في علم البيانات باستخدامPython (ساعتان).

الهدف: توفير أساس متين لعلم البيانات باستخدام Python من خلال التعرف على المكتبات الأساسية المستخدمة في علم البيانات، فهم هيكلة البيانات وإتقان الأوامر الأساسية اللازمة لاستكشاف البيانات وفحصها.

الأدوات: Jupyter Notebook، Google Colab، Python 3.8+، مكتبة Pandas، مكتبة NumPy.

المحتوى:

- تحميل البيانات (CSV، Excel، JSON)

```
import pandas as pd
df_csv = pd.read_csv('data.csv')
df_excel = pd.read_excel('data.xlsx')
df_json = pd.read_json('data.json')
```

- نظرة عامة على إطار البيانات (DataFrame)

```
df.head()                    # أول 5 صفوف
df.tail()                    # آخر 5 صفوف
df.sample(3)                 # 3 صفوف عشوائية
df.shape                     # الأبعاد (عدد الصفوف، عدد الأعمدة)
df.columns                   # أسماء الأعمدة
df.dtypes                    # نواع كل عمود
df.info()                    # ملخص كامل
```

- الإحصائيات الأساسية

```
df.describe()                # ملخص إحصائي رقمي
df['colonne'].value_counts() # تكرار القيم
df['colonne'].nunique()      # عدد القيم الفريدة
```

- تكوين العرض

```
pd.set_option('display.max_columns', None)  # عرض جميع الأعمدة
pd.set_option('display.precision', 2)       # عدد الأرقام العشرية
```

- اختيار أعمدة/صفوف

```
df['marque']                 # عمود واحد
df[['marque', 'prix']]       # عدة أعمدة
df.iloc[0]                   # السطر الأول
df.iloc[0:5]                 # الأسطر من 0 إلى 5
df.loc[df['prix'] > 10000]   # تصفية حسب الشرط
```

- إضافة أعمدة وحذفها

```
df['prix_par_kg'] = df['prix'] / df['poids']
df.drop('ancien_colonne', axis=1, inplace=True)
```

- فرز (ترتيب) قيم

```
df.sort_values(by='prix', ascending=False)
```

- **ورشة عمل تطبيقية:**

  - تحميل ملف CSV يحتوي على معلومات عن المركبات
  - عرض الأعمدة، الأنواع، والأبعاد
  - حساب القيم الفريدة في عمود (مثال: الوقود)
  - إنشاء عمود محسوب (السعر_لكل_كيلوغرام)
  - ترتيب السيارات حسب السعر تنازليًا واستخراج أول 10 سيارات

2- تحضير البيانات وتنقيتها (3 ساعات)

الهدف: تحسين جودة البيانات وضمان اتساقها وقوتها من خلال تنظيفها وتصحيحها وتحويلها، مما يساهم في تعزيز دقة التحليلات والنماذج التنبؤية.

الأدوات: Pandas، NumPy، Scikit-learn (لإعداد البيانات)، OpenRefine (اختياري).

المحتوى:

- كشف القيم المفقودة

  - تحديد أماكن القيم المفقودة NaN حسب العمود أو الصف

```
df.isnull().sum()
df[df.isnull().any(axis=1)]
```

- معالجة القيم المفقودة

  - التعبئة (التضمين) حسب المتوسط أو الوسيط أو الثابت أو الطريقة المتقدمة

```
df['age'] = df['age'].fillna(df['age'].mean())
df['sexe'] = df['sexe'].fillna('Inconnu')
df = df.ffill()                    # ملء القيم المفقودة بطريقة التوسيع للأمام
df = df.bfill()                    # ملء القيم المفقودة بطريقة التوسيع للخلف
```

  - الاستكمال باستخدام SimpleImputer (Scikit-learn)

```
from sklearn.impute import SimpleImputer
imputer = SimpleImputer(strategy='median')
df[['revenu']] = imputer.fit_transform(df[['revenu']])
```

- الكشف عن القيم الشاذة (القيم المتطرفة)

  - طريقة الفارق بين الربيعيات(IQR)

```
Q1 = df['prix'].quantile(0.25)
Q3 = df['prix'].quantile(0.75)
IQR = Q3 - Q1
outliers = df[(df['prix'] < Q1 - 1.5 * IQR) | (df['prix'] > Q3 + 1.5 * IQR)]
```

- التصفية حسب درجة z (z-score)

```
from scipy.stats import zscore
df['zscore'] = zscore(df['prix'])
df[df['zscore'].abs() > 3]
```

- حذف التكرارات (البحث وإزالة الأسطر المكررة)

```
df.duplicated().sum()
df.drop_duplicates(inplace=True)
```

- تنقية سلاسل النصوص

```
df['ville'] = df['ville'].str.strip().str.lower().str.replace('-', ' ')
df['ville'] = df['ville'].str.normalize('NFKD')
```

- توحيد أنواع البيانات

```
df['date'] = pd.to_datetime(df['date'])
df['prix'] = df['prix'].astype(float)
```

- إنشاء أعمدة مشتقة

- إنشاء متغيرات جديدة مفيدة استنادًا إلى المتغيرات الحالية

```
df['prix_par_kg'] = df['prix'] / df['poids']
df['anciennete'] = 2024 - df['annee']
```

- ترميز المتغيرات الفئوية

- الترميز الأحادي وترميز التسميات

```
pd.get_dummies(df, columns=['carburant'])
from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
df['sexe_code'] = le.fit_transform(df['sexe'])
```

- التطبيع وتوحيد المقاييس

- قياس النماذج الحساسة (KNN، الانحدار الخطي، إلخ)

```
from sklearn.preprocessing import MinMaxScaler, StandardScaler
MinMaxScaler().fit_transform(df[['revenu']])
StandardScaler().fit_transform(df[['revenu']])
```

- الكشف عن التناقضات في العمل (مثال :وزن أقل من 300 كجم أو عمر أكبر من 120 سنة)

```
df[df['poids'] < 300]
df[df['age'] > 120]
```

- **ورشة عمل تطبيقية:**

  - تحديد القيم المفقودة والشاذة وغير المتسقة ومعالجتها.
  - تنظيف الأعمدة النصية وتوحيد التنسيقات.
  - إنشاء أعمدة مفيدة للنمذجة.
  - تطبيق التطبيع وترميز الفئات.

3- التجميع والوصلات والتجميعات (ساعتان)

الهدف: تنظيم البيانات وتلخيصها على عدة محاور: نوع المركبة، والوقود، والفترة الزمنية، وما إلى ذلك. الحصول على مؤشرات الأداء الرئيسية، وإنشاء جداول مرجعية متقاطعة ومقارنة عدة مصادر.

الأدوات: Pandas.

المحتوى:

- حساب المتوسطات، المجاميع، وعدد العناصر لكل فئة

```
df.groupby('marque')[ df].mean()
df.groupby('carburant')['puissance'].sum()
df.groupby('type').size()                                    # حساب عدد العناصر لكل نوع
```

- التجميعات المتعددة باستخدام الأمر ()agg

```
df.groupby('marque').agg({
  'prix': ['mean', 'max', 'std'],
  'poids': 'median'
})
```

- التجميع على عدة أعمدة في وقت واحد

```
gb = df.groupby(['marque', 'carburant'])
gb['prix'].mean()
```

- الجداول المحورية الديناميكية (Pivot Tables)

```
pd.pivot_table(df, values='prix', index='marque', columns='carburant', aggfunc='mean')
```

- فرز (ترتيب) المجموعات المجمَّعة

```
df.groupby('marque')['prix'].mean().sort_values(ascending=False)
```

- دمج إطارات البيانات DataFrames (الوصلات العلائقية)

```
pd.merge(df_clients, df_commandes, on='client_id', how='inner')
pd.merge(df1, df2, left_on='id', right_on='produit_id', how='outer')
```

- دمج البيانات (تكديس عدة جداول)

```
pd.concat([df1, df2], axis=0)                    # دمج البيانات صفاً بصف
pd.concat([df1, df2], axis=1)                    # دمج البيانات عموداً بعمود
```

- **ورشة عمل تطبيقية:**
  - إنشاء جدول محورية ديناميكي لمتوسط الأسعار حسب العلامة التجارية ونوع الوقود.
  - دمج بيانات العملاء وبيانات الطلبات >>لحساب إجمالي الإنفاق لكل عميل.

4- التحليل الاستكشافي والإحصائيات (ساعتان)

الهدف: اكتساب الأساسيات الإحصائية لفهم أفضل لمتغيرات مجموعة بيانات، بما في ذلك الاتجاهات المركزية، التشتت، التباين، الارتباط البسيط، والتصور الواضح.

الأدوات: Pandas, NumPy, Seaborn, Matplotlib, Statsmodels

المحتوى:

- والإحصائيات الأساسية
  - المتوسط، الوسيط، الانحراف المعياري والرباعيات

```
df.describe()
df['prix'].mean(), df['prix'].std(), df['prix'].quantile([0.25, 0.5, 0.75])
```

- التباين والتغاير

```
df.var()
df.cov()
df.corr()                                    # حساب معامل الارتباط بيرسون بين الأعمدة
```

- عرض التوزيعات
  - الرسم البياني التكراري والرسم البياني الصندوقي

```
import seaborn as sns
sns.histplot(df['prix'], bins=20, kde=True)
sns.boxplot(x=df['prix'])
```

- تحليل التباين بين المجموعات (ANOVA مبسط)

```
from scipy.stats import f_oneway
f_oneway(df[df['marque'] == 'Peugeot']['prix'], df[df['marque'] == 'Renault']['prix'])
```

- الانحدار الخطي البسيط باستخدام Statsmodels

```
import statsmodels.api as sm
X = df[['poids']]
y = df['prix']
X = sm.add_constant(X)
model = sm.OLS(y, X).fit()
print(model.summary())
```

- **ورشة عمل تطبيقية :**
  - إجراء مقارنة بصرية لتوزيع الأسعار حسب العلامات التجارية.
  - إظهار الإحصائيات الوصفية لمتغير عددي.
  - تحليل الانحراف المعياري والقيم الربعية بين المجموعات.

5- عرض البيانات (3 ساعات)

الهدف: تقديم العلاقات بين المتغيرات بشكل مرئي لتبسيط التحليل والتفسير.

الأدوات: Matplotlib, Seaborn, Plotly

- المدرج التكراري

```
import seaborn as sns
sns.histplot(df['prix'], bins=30)
```

- سحابة نقطية

```
sns.scatterplot(x='poids', y='prix', hue='marque', data=df)
```

- خريطة الارتباطات الحرارية

```
sns.heatmap(df.corr(), annot=True, cmap='coolwarm')
```

- عروض تفاعلية باستخدام Plotly

```
import plotly.express as px
px.scatter(df, x='poids', y='prix', color='carburant')
```

- **ورشة عمل تطبيقية:**
  - إجراء مقارنة بصرية لطرازات السيارات حسب أدائها وأسعارها.

6- النمذجة التنبؤية والتعلم العميق (10 ساعات)

الهدف: بناء نماذج تنبؤية ومقارنتها، باستخدام النماذج الكلاسيكية والنماذج البسيطة في التعلم العميق.

الأدوات: Scikit-learn, XGBoost, Keras, SHAP, Joblib, pycaret

هيكل الوحدة (10 ساعات)

- **ساعة:** إعداد البيانات وتقسيمها إلى مجموعات تدريب واختبار.
- **ساعتان:** تطبيق الانحدار الخطي والانحدار المنتظم Lasso وRidge.
- **ساعتان:** التصنيف باستخدام تقنيات Logistic Regression، SVM، KNN.
- **ساعتان:** استكشاف طرق متقدمة مثل Random Forest، XGBoost، LGBM.
- **ساعة:** التعرف على أساسيات التعلم العميق باستخدام Keras والشبكات الكثيفة.
- **ساعتان:** التقييم المتقدم للنماذج، تفسير النتائج باستخدامSHAP ، وحفظ النماذج.

المحتوى:

- إعداد مجموعة البيانات

```
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
```

- نماذج الانحدار

```
from sklearn.linear_model import LinearRegression
model = LinearRegression()
model.fit(X_train, y_train)
```

- الانحدار باستخدام الغابة العشوائية

```
from sklearn.ensemble import RandomForestRegressor
rf = RandomForestRegressor()
rf.fit(X_train, y_train)
```

- الانحدار باستخدام التعزيز

```
from xgboost import XGBRegressor
xgb = XGBRegressor()
xgb.fit(X_train, y_train)
```

- نماذج التصنيف

```
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import classification_report
clf = LogisticRegression()
clf.fit(X_train, y_train)
y_pred = clf.predict(X_test)
print(classification_report(y_test, y_pred))
```

- الشبكات العصبية البسيطة باستخدام Keras

```
from keras.models import Sequential
from keras.layers import Dense
model = Sequential()
model.add(Dense(10, input_dim=X_train.shape[1], activation='relu'))
model.add(Dense(1, activation='linear'))
model.compile(optimizer='adam', loss='mse')
model.fit(X_train, y_train, epochs=100, batch_size=10)
```

- التقييم والتفسير

```
from sklearn.metrics import mean_squared_error, r2_score
print(mean_squared_error(y_test, y_pred))
print(r2_score(y_test, y_pred))
```

- تفسير النتائج باستخدام SHAP

```
import shap
explainer = shap.Explainer(model.predict, X_test)
shap_values = explainer(X_test)
shap.plots.beeswarm(shap_values)
```

- استخدام AutoML مع PyCaret

- **PyCaret** هي مكتبة مفتوحة المصدر بلغة Python تسهل تطوير نماذج التعلم الآلي. تعتبر بديلاً مبسطًا لـ scikit-learn، حيث تم تصميمها لتسريع وأتمتة دورة الحياة الكاملة للنمذجة: من تحضير البيانات إلى نشر النماذج في بيئة إنتاجية.
- واجهة بسيطة للغاية :بضع أسطر من التعليمات البرمجية تكفي.
- معالجة تلقائية للبيانات: مثل التعامل مع القيم المفقودة، الترميز، والتطبيع.
- مقارنة تلقائية بين العديد من الخوارزميات.

- تدعم العديد من حالات الاستخدام:
  - التصنيف (Classification)
  - الانحدار (Regression)
  - التجميع (Clustering)
  - اكتشاف الحالات الشاذة (Anomaly Detection)
  - تحليل النصوص(NLP)
  - السلاسل الزمنية(Forecasting)

مثال: تطبيق الانحدار التلقائي على مجموعة بيانات السيارات

```
# استيراد المكتبات
from pycaret.regression import *
import pandas as pd
# تحميل مجموعة البيانات
df = pd.read_csv("voitures.csv")
# AutoML تهيئة بيئة
reg = setup(data=df, target='prix', session_id=123, verbose=True)
# مقارنة تلقائية بين جميع النماذج
best_model = compare_models()
# عرض المخلفات وأهمية المتغيرات
plot_model(best_model, plot='residuals')
# عرض أهمية المتغيرات
plot_model (best_model, plot='feature')
# حفظ النموذج
save_model(best_model, 'modele_auto_prix')
```

- **ورشة عمل تطبيقية:**
  - مقارنة النماذج: الانحدار الخطي مقابل الغابة العشوائية مقابل الشبكة العصبية
  - اختياري: إعادة تطبيق بعض نماذج التعلم الآلي باستخدام PyCaret.

▪ استخدام AutoML مع h2o.ai

- h2o.ai هي مكتبة مفتوحة المصدر بلغة R، تساهم في تبسيط وتسريع تطوير نماذج التعلم الآلي. تعد بديلا قويا لأتمتة الدورة الكاملة للنمذجة، بدءًا من تحضير البيانات وصولًا إلى نشر النماذج في بيئة إنتاجية.
- واجهة سهلة الاستخدام: تتيح h2o.ai تطوير نماذج معقدة باستخدام بضعة أسطر من التعليمات البرمجية فقط.
- المعالجة التلقائية للبيانات: تشمل التعامل مع القيم المفقودة، وترميز المتغيرات الفئوية، والتطبيع، وغير ذلك.
- المقارنة التلقائية بين عدة خوارزميات: تتيح اختبار خوارزميات تعلم آلي مختلفة لاختيار الأكثر كفاءة.
- يدعم العديد من حالات الاستخدام، منها:
  - التصنيف (Classification)
  - الانحدار (Regression)
  - التجميع (Clustering)

- اكتشاف الحالات الشاذة (Anomaly Detection)
- تحليل النصوص(NLP)
- السلاسل الزمنية(Forecasting)

```
# تحميل مكتبة h2o
library(h2o)

# تهيئة خادم H2O
h2o.init()

# تحميل مجموعة البيانات
df <- h2o.importFile("voitures.csv")                    # يحتوي على الوزن، القوة، السعر، إلخ

# تهيئة بيئة AutoML
aml <- h2o.automl(y = "prix", training_frame = df, max_models = 10, seed = 1234)

# المقارنة التلقائية لجميع النماذج
leaderboard <- aml@leaderboard
print(leaderboard)

# عرض أداء النموذج الأفضل
best_model <- aml@leader
performance <- h2o.performance(best_model, newdata = df)
print(performance)

# حفظ النموذج
h2o.saveModel(best_model, path = "modele_auto_prix", force = TRUE)
```

- **ورشة عمل تطبيقية:**

- مقارنة النماذج: الانحدار الخطي مقابل الغابة العشوائية مقابل الشبكة العصبية باستخدام h2o.ai
- اختياري: إعادة تجربة بعض نماذج التعلم الآلي باستخدام h2o.ai لمراقبة الأداء على مجموعات بيانات حقيقية، مثل مجموعة بيانات المركبات.

7- إنشاء لوحة تحكم تفاعلية (ساعتان)

الهدف: إنشاء تطبيق تفاعلي بسيط باستخدام Python فقط والمكتبات القياسية لعرض وتحليل البيانات. بالإضافة إلى دمج TensorBoard لعرض وتتبع مقاييس النماذج في التعلم الآلي.

الأدوات: Pandas, Matplotlib, ipywidgets, TensorBoard (via TensorFlow)

مثال تفصيلي:

- عرض تفاعلي لأسعار السيارات حسب الوزن

```
import pandas as pd
import matplotlib.pyplot as plt
from ipywidgets import interact, IntSlider
```

- تحميل مجموعة بيانات كمثال

```
data = {'poids': [900, 1200, 1500, 1800, 2000], 'prix': [7000, 12000, 18000, 25000, 30000]}
df_voitures = pd.DataFrame(data)
```

- وظيفة تفاعلية لعرض البيانات

```
def visualiser_prix_max(poids_max):
df_filtre = df_voitures[df_voitures['poids'] <= poids_max]
plt.figure(figsize=(8, 4))
plt.scatter(df_filtre['poids'], df_filtre['prix'], color='blue')
plt.title(f'Prix des voitures jusqu'à {poids_max} kg')
plt.xlabel('Poids (kg)')
plt.ylabel('Prix (€)')
plt.grid(True)
plt.show()
```

- عنصر تحكم الواجهة الرسومية التفاعلية

```
interact(visualiser_prix_max, poids_max=IntSlider(min=800, max=2000, step=100, value=1500))
```

- استخدام TensorBoard لتتبع وعرض عملية تدريب نموذج بسيط

```
import tensorflow as tf
import numpy as np
# بيانات بسيطة
X = np.array([[0], [1], [2], [3], [4]], dtype=float)
y = np.array([[0], [1], [4], [9], [16]], dtype=float)
# إنشاء نموذج بسيط
model = tf.keras.Sequential([
tf.keras.layers.Dense(units=10, activation='relu', input_shape=[1]),
tf.keras.layers.Dense(units=1)
])
model.compile(optimizer='adam', loss='mse')
log_dir = "logs/fit"
tensorboard_callback = tf.keras.callbacks.TensorBoard(log_dir=log_dir, histogram_freq=1)
model.fit(X, y, epochs=50, callbacks=[tensorboard_callback])
```

- استخدام TensorBoard لعرض النتائج داخل دفتر Jupyter أو Colab

```
%load_ext tensorboard
%tensorboard --logdir logs/fit
```

- **ورشة عمل تطبيقية :**

- قم بتنزيل وتحميل ملف CSV يحتوي على بيانات حقيقية عن المركبات.
- قم بعرض الإحصائيات الوصفية لتحليل البيانات المحملة.
- عدّل الوظيفة التفاعلية السابقة لعرض الأسعار بناءً على معيار مختلف، مثل القوة الحصانية أو سنة الإنتاج.
- استكشف واستخدم أدوات واجهة تفاعلية إضافية لتحسين تجربة المستخدم.

- اختياريًا: قم بتدريب نموذج بسيط واعرض نتائجه باستخدام أداة TensorBoard

المحور الرابع: توظيف الذكاء الاصطناعي في حل المشكلات البحثية (8 ساعات)

يهدف هذا المحور إلى مساعدة طلبة الدكتوراه على تطبيق الذكاء الاصطناعي من خلال دراسات حالة مخصصة لمجالات علمية مختلفة. سيتعلم الطلبة إتقان سلسلة تطوير عملية الذكاء الاصطناعي، بدءًا من معالجة البيانات الأولية وصولاً إلى تحليلها وإنتاج النتائج، مع الاعتماد على أدوات متخصصة. سيتم التركيز على التجريب باستخدام قواعد بيانات حقيقية، بالإضافة إلى إنشاء لوحات تحكم تفاعلية لتصور النتائج وتفسيرها بشكل واضح وفعّال. سيتم تطبيق هذه المعارف من خلال ورش عمل مخصصة، تتيح لطلبة الدكتوراه فرصة تجربة التقنيات المطروحة عمليًا.

1- إنشاء لوحات معلومات تفاعلية لتحليل البيانات (ساعتان)

الهدف: تعلم كيفية إنشاء لوحات تحكم تفاعلية وديناميكية لتحليل وعرض البيانات العلمية بشكل فعال. تتيح هذه الأدوات استكشافا عميقا للبيانات وتسهل التواصل الواضح للنتائج، مما يجعل المعلومات المعقدة أكثر سهولة وفهما.

الأدوات: Power BI (ou Tableau), Shiny (R)

المحتوى:

- مقدمة عن لوحات التحكم
  - تقديم الأدوات Power BI، Tableau، Shiny
  - المبادئ الأساسية لعرض البيانات.
  - هيكلة لوحة التحكم (الأهداف، المؤشرات الرئيسية، وتأثيرها على البحث).
- إنشاء لوحة تحكم
  - استيراد البيانات وتحضيرها.
  - إنشاء تصورات بسيطة (الرسوم البيانية، منحنيات الاتجاه).
  - دمج الوظائف التفاعلية (المرشحات، الاستكشاف المتدرج)
- التحليل الاستكشافي
  - استكشاف البيانات بشكل تفاعلي.
  - تحديد الاتجاهات أو الحالات الشاذة.

- **ورشة عمل تطبيقية :**

- يعمل طلبة الدكتوراه على تصميم لوحة معلومات تفاعلية باستخدام مجموعة بيانات علمية (مثل البيانات البيئية أو الطبية).
- يقومون بتحليل البيانات لاستكشاف الاتجاهات العامة والكشف عن الحالات الشاذة أو الأنماط غير المعتادة.

2- دراسات حالة: تطبيق الذكاء الاصطناعي في مجالات متنوعة (3 ساعات)

الهدف: استخدام تقنيات الذكاء الاصطناعي لمعالجة مشكلات حقيقية في مختلف المجالات. من خلال تطبيق الذكاء الاصطناعي على تحديات معينة، يتم تسليط الضوء على فائدته وإمكاناته في تحويل البحث العلمي، من خلال تقديم حلول مبتكرة وفتح آفاق جديدة لحل المشكلات المعقدة.

الأدوات: حسب المجال (العلوم الإنسانية والاجتماعية أو العلوم والتكنولوجيا) والمشروع المختار.

المحتوى:

- تقديم حالات استخدام الذكاء الاصطناعي في مجالات متنوعة.
- شرح تقنيات الذكاء الاصطناعي المناسبة لكل مجال دراسي.
- تنفيذ نموذج ذكاء اصطناعي مناسب للمشكلة المختارة.
- تحليل النتائج وتقديم مقترحات لتحسينها.
- مناقشة القضايا والتحديات الخاصة بتطبيق الذكاء الاصطناعي في المجال.
- دراسة نقدية: نقاط القوة والضعف والقيود التي يواجهها الذكاء الاصطناعي في البحث الأكاديمي.

أمثلة لمشاريع في العلوم الإنسانية والاجتماعية

- **الهدف:** المسح الضوئي وتحليل الوثائق القديمة (مثل الرسائل، الصحف) لاستخراج معلومات رئيسية (التواريخ، الأماكن، الأشخاص...).
- **التقنيات:** التعرف الضوئي على الحروف (OCR)، التعرف على الكيانات المسماة (NER)، تصنيف النصوص.
- **الأدوات:** Tesseract (OCR)، SpaCy، Python.

مثال لمشروع في العلوم والتكنولوجيا

- **الهدف:** تصنيف الصور الطبية (مثل الأشعة السينية، التصوير بالرنين المغناطيسي) لاكتشاف الأمراض.
- **التقنيات:** الشبكات العصبية التلافيفية (CNN)، التعلم المنتقل (transfer learning).
- **الأدوات:** Keras، TensorFlow، OpenCV.

مثال لمشاريع متعددة التخصصات:

- **الهدف:** تطوير روبوت محادثة (chatbot) قادر على الإجابة على الأسئلة في مجالات العلوم الإنسانية والاجتماعية والعلوم والتكنولوجيا (مثل الأسئلة حول تاريخ العلوم والمفاهيم التقنية).
- **التقنيات:** نماذج اللغةGPT ، BERT، معالجة اللغة الطبيعية NLP.
- **الأدوات:** Rasa، Dialogflow، Transformers.

- **ورشة عمل تطبيقية :**

- **العمل الجماعي على دراسة حالة:** سيقوم طلبة الدكتوراه بتطبيق تقنيات الذكاء الاصطناعي على مشكلة واقعية.
- **تنفيذ نموذج:** سيقوم طلبة الدكتوراه بتدريب وتقييم نموذج للذكاء الاصطناعي.
- **تحليل النتائج:** سيقوم طلبة الدكتوراه بتفسير النتائج المستخلصة وتقديم اقتراحات للتحسين.

3- التجريب على قواعد بيانات علمية (ساعتان)

الهدف: توفير تجربة عملية لطلبة الدكتوراه من خلال العمل على مجموعات بيانات علمية حقيقية. سيقوم الطلاب بالاستكشاف، المعالجة، والتحليل لهذه البيانات، مع مواجهة التحديات الخاصة بكل مجال. من خلال هذه المقاربة، سيتعلمون كيفية التعامل مع مجموعات بيانات متنوعة، واستخراج المعلومات المهمة، وتطوير مهاراتهم في إدارة وتحليل البيانات. سيمكنهم هذا الجزء من التكوين من التكيف مع البيانات المعقدة وفهم أفضل لاستخدامها في سياق البحث العلمي.

الأدوات: Python، Pandas، NumPy، Matplotlib/Seaborn، Scikit-learn، Jupyter Notebook

المحتوى:

- مقدمة حول مجموعات البيانات العلمية
  - تعريف أنواع مجموعات البيانات العلمية (البيانات الجدولية، السلاسل الزمنية، الصور، النصوص، إلخ).
  - التحديات الخاصة المرتبطة بمجموعات البيانات العلمية (جودة البيانات، الحجم، التفاوت، التحيز).
  - أمثلة على مجموعات بيانات علمية عامة مثل UCI Machine Learning Repository، Kaggle، بوابات البيانات المفتوحة.
- استكشاف وتحضير البيانات
  - تقنيات تنظيف البيانات (إدارة القيم المفقودة، التقييس، التوحيد).
  - استكشاف البيانات (الإحصائيات الوصفية، عرض التوزيعات).
  - إعداد البيانات للتحليل (هندسة الميزات، اختيار المتغيرات).
- تحليل البيانات
  - تطبيق تقنيات التحليل الإحصائي (الارتباطات، اختبارات الفرضيات).
  - استخدام نماذج تعلم الآلة البسيطة لاستكشاف البيانات (مثل: التجميع، الانحدار الخطي).
  - عرض النتائج (الرسوم البيانية، البطاقات، إلخ).
- دراسة نقدية
  - مناقشة حدود مجموعات البيانات العلمية (جودة البيانات، التحيزات المحتملة، التمثيلية).
  - التفكير في تأثير الخيارات المنهجية على نتائج البحث.

5. المراجع العلمية

1. Ian Goodfellow, Yoshua Bengio, Aaron Courville – *Deep Learning*, MIT Press.
2. Aurélien Géron – *Hands-on Machine Learning with Scikit-Learn, Keras & TensorFlow*, O'Reilly.
3. Stuart Russell, Peter Norvig – *Artificial Intelligence: A Modern Approach*, Pearson.
4. Ian Goodfellow – *GANs (Generative Adversarial Networks)*, MIT Press.
5. Christopher M. Bishop – *Pattern Recognition and Machine Learning*, Springer.
6. François Chollet – *Deep Learning with Python*, Manning.
7. Sebastian Raschka, Vahid Mirjalili – *Python Machine Learning*, Packt Publishing.
8. Kevin P. Murphy – *Machine Learning: A Probabilistic Perspective*, MIT Press.
9. Daphne Koller, Nir Friedman – *Probabilistic Graphical Models: Principles and Techniques*, MIT Press.
10. Tom Mitchell – *Machine Learning*, McGraw-Hill.
11. Jure Leskovec, Anand Rajaraman, Jeffrey Ullman – *Mining of Massive Datasets*, Cambridge University Press.
12. Michael Nielsen – *Neural Networks and Deep Learning* (disponible gratuitement en ligne).
15. Fast.ai – Cours gratuits sur l'IA et le Deep Learning.
16. Coursera - Deep Learning Specialization – Andrew Ng, Stanford.
17. Udacity - AI for Everyone – Cours introductif sur l'IA.
18. MIT OpenCourseWare – Cours IA du MIT.
19. Deep Learning AI – Plateforme d'apprentissage dirigée par Andrew Ng.
20. Stanford CS231n – Cours sur les réseaux de neurones convolutifs pour la vision par ordinateur.
21. Google Colab – Plateforme pour exécuter des modèles IA gratuitement en cloud.
22. Kaggle – Compétitions et tutoriels pratiques en Machine Learning.
23. Hugging Face Courses – Cours spécialisés sur le NLP et les Transformers.
24. Google AI – Ressources et cours de Google sur l'IA et le ML.
25. OpenAI Blog – Articles et publications sur les dernières avancées en IA.
26. PyTorch Tutorials – Tutoriels pour apprendre PyTorch en profondeur.
27. Pandas Documentation – Documentation officielle pour l'analyse des données avec Python.
28. Seaborn Documentation – Visualisation statistique avancée.
29. Scikit-Learn Guide – Ressource pour la modélisation prédictive en Python.
30. Dash for Beginners: Create Interactive Data Apps with Plotly and Dash.